Backtracking Bipolar Magnetic Regions to Their Emergence: Two Groups and Their Implication in the Tilt Measurements

Anu Sreedevi[1] , Bidya Binay Karak[1] , Bibhuti Kumar Jha[2] , Rambahadur Gupta[1], and Dipankar Banerjee[3,4]
[1] Department of Physics, Indian Institute of Technology (Banaras Hindu University), Varanasi, 221005, India; anubsreedevi.rs.phy20@itbhu.ac.in,
karak.phy@iitbhu.ac.in
[2] Southwest Research Institute, Boulder, CO 80302, USA
[3] Indian Institute of Astrophysics, Koramangala, Bangalore 560034, India
[4] Center of Excellence in Space Sciences India, IISER Kolkata, Mohanpur 741246, West Bengal, India
Received 2025 August 30; revised 2025 November 7; accepted 2025 November 20; published 2025 December 31

Abstract

Bipolar magnetic regions (BMRs) that appear on the solar photosphere are surface manifestations of the Sun's internal magnetic field. With modern observations and continuous data streams, the study of BMRs has moved from manual sunspot catalogs to automated detection and tracking methods. In this work, we present an additional module to the existing BMR tracking algorithm, the Automatic Tracking Algorithm for Bipolar Magnetic Regions (AutoTAB), which focuses on identifying emerging signatures of BMRs. Specifically, for regions newly detected on the solar disk, this module backtracks the BMRs to their point of emergence. From a total of about 12,000 BMRs identified by AutoTAB, we successfully backtracked 3080 cases. Within this backtracked sample, we find two distinct populations. One group shows the expected behavior of emerging regions, in which the magnetic flux increases significantly during the emerging phase. The other group consists of BMRs whose flux, however, does not exhibit substantial growth during their evolution, the instances where our algorithm fails to capture the initial emergence of the BMRs. We classify these as "discarded" BMRs and examine their statistical properties separately. Our analysis shows that these discarded BMRs do not display any preferred tilt angle distribution and do not show systematic latitudinal tilt dependence, in contrast to the trends typically associated with emerging BMRs. This indicates that including such regions in statistical studies of BMR properties can distort or mask the underlying physical characteristics. We therefore emphasise the importance of excluding the discarded population from the whole dataset when analysing the statistical behavior of BMRs.

Unified Astronomy Thesaurus concepts: Sunspots (1653); Solar active region magnetic fields (1975); Solar activity (1475); Solar cycle (1487); Sunspot cycle (1650); Bipolar sunspot groups (156); Sunspot groups (1651); Solar physics (1476); Solar interior (1500); Solar surface (1527); Solar magnetic fields (1503)

1. Introduction

Isolated, intense magnetic field regions on the solar photosphere manifest as dark spots, often appearing in pairs or groups, in white-light images. When observed in magnetically sensitive spectral lines, these sunspot pairs correspond to opposite polarities of a more general magnetic structure known as bipolar magnetic regions (BMRs; G. E. Hale 1908; R. F. Howard 1991a, 1991b). Continuous records of sunspots have been maintained since the early 1600s, following the popularization of telescopes, while systematic observations of their magnetic counterparts date back to the late 1900s with the advent of magnetographs (K. R. Sivaraman et al. 1999). Owing to the long-standing observational record, these features have been studied for over a century. Since sunspots are the primary sites of solar eruptive events, they serve as a key indicator of solar activity (C. J. Schrijver 2009). Consequently, study of the evolution and properties of individual BMRs is essential for understanding and improving the predictability of the solar events and solar cycle (K. Petrovay 2020; B. B. Karak 2023).

The cataloging of solar magnetic regions has traditionally relied on human observers and sunspot records. In the last

decade, data from various ground-based observatories have been digitized and made publicly available, leading to the development of several comprehensive, cross-calibrated sunspot catalogs (R. Arlt & J. M. Vaquero 2020). The advent of high-cadence satellite observations in recent years—Michelson Doppler Imager (MDI) on board Solar and Heliospheric Observatory (SOHO) and the Helioseismic and Magnetic Imager (HMI) on board Solar Dynamic Observatory (SDO)—demands a better way of studying and cataloging BMRs out of the reach of human observers. This helped in the development of automatic techniques to detect and keep track of the evolving BMRs and sunspots through their nearside and farside evolution. Starting initially with automatic detection, A. G. Kosovichev & J. O. Stenflo (2008) and J. O. Stenflo & A. G. Kosovichev (2012) introduced a method for automatically isolating strictly bipolar structures in line-of-sight (LOS) magnetograms through a flux-based threshold, a technique later adapted to HMI data by B. K. Jha et al. (2020) and A. Sreedevi et al. (2023). Subsequently, A. G. Tlatov & A. A. Pevtsov (2014) developed an automated algorithm for identifying penumbrae and umbrae from daily HMI white-light images. Threshold-based approaches have since improved through morphological operations such as opening/closing and region growth techniques. However, analyses based on such datasets tend to emphasize long-lived BMRs, as they persist longer in the dataset. This bias can be mitigated by tracking BMRs throughout their evolution, allowing

researchers to represent a given region at any desired stage and led to development of various tracking algorithms of BMRs. The Solar Monitor Active Region Tracking (P. A. Higgins et al. 2011) algorithm became one of the early tools for real-time detection and tracking, aimed at supporting solar eruptive event prediction. Optimized for space weather prediction, they overlook small features. The Bipolar Active Region Detection (A. Muñoz-Jaramillo et al. 2016) system and dataset from R. Wang et al. (2024) employ a similar detection framework but employ a dual-maximum flux-weighted overlap method for feature association and incorporate human oversight for complicated pairing to ensure efficiency. Similar datasets are provided by Space-Weather MDI Active Region Patches (SMARPs; M. G. Bobra et al. 2014) and Space-Weather HMI Active Region Patches (SHARPs; M. G. Bobra et al. 2021), which extend the tracking capability by providing magnetogram-based maps of active regions identified in MDI and HMI observations. However, active regions in SMARPs and SHARPs are not necessarily true BMRs, as they may lack sufficient flux balance. More recently, the rapid rise of machine learning has brought artificial intelligence (AI)–based detection and tracking models into widespread use, offering speed and adaptability beyond traditional methods (L. Gong et al. 2024). Yet a fundamental challenge remains: most existing techniques, whether manual, threshold-based, tracking-based, or AI-driven, still fail to capture the very early stages of BMR emergence and rarely differentiate between the emerging BMRs and the nonemerging phase of the BMRs or the bipolar structures arising from fragmentation of decaying BMRs.

In this study, we aim to identify and track BMRs from their earliest stages of emergence, distinguishing BMRs from the emerging/decaying phase and fragmented regions and examining their properties. We shall show that a significant fraction of BMRs detected and tracked in the LOS magnetograms do not correspond to true emerging phases of BMRs, and their properties (e.g., the tilt angle) are quite different from what one expects for true emerging BMRs. We shall show that it is essential to isolate these regions from the primary data before studying the properties of true emerging BMRs.

The paper is organized as follows. Sections 2 and 3, respectively, describe the datasets used and the working of our automatic detection algorithm. Section 4 details the methodology for classifying BMRs as truly emerging regions and the others and analyzes their respective properties. Finally, Section 5 summarizes our findings and presents concluding remarks.

2. Data

For our study we use the LOS magnetogram data from two instruments, namely, MDI (P. H. Scherrer et al. 1995) on board SOHO and HMI (P. H. Scherrer et al. 2012) on board SDO. MDI provides LOS magnetograms with a spatial resolution of approximately $4''$ and a cadence of 96 minutes from 1996 to 2011, whereas HMI provides the same with an exceptional spatial and temporal resolution of $1''$ every 45 s and 12 minutes since 2010. The HMI LOS magnetograms are rebinned to the MDI resolution and scaled by a factor of 1.4 for consistency. In this study, we use the full dataset spanning 1996–2024.

3. Methodology

For the objective mentioned in the Introduction, we employ our algorithm in the LOS magnetograms. The algorithm itself operates in three stages: (1) automatic detection, (2) automatic tracking, and (3) automatic backtracking to identify emergence. The first two stages are implemented as part of the Automatic Tracking Algorithm for Bipolar Magnetic Regions (AutoTAB; A. Sreedevi et al. 2023). Although the details of AutoTAB are provided in the referenced article, the algorithm has since undergone several improvements. Therefore, we start with a brief description of AutoTAB, followed by the backtracking procedure.

3.1. AutoTAB

AutoTAB is an automated state-of-the-art tool designed to track the evolution of BMRs during its passage across the visible solar disk as seen from the Sun–Earth line. The algorithm constitutes two major steps: the detection and tracking of BMRs.

We begin by correcting the LOS magnetic field measurements for projection effects. This correction is done by dividing the magnetic field value at each pixel by the cosine of its heliocentric angle, the angle between the LOS, and the normal to the solar surface at that pixel's location. Next, we smooth the image using a boxcar kernel of the size of 11 pixels ($22''$) and apply an adaptive threshold based on the average magnetic field strength in the image (J. O. Stenflo & A. G. Kosovichev 2012). This step helps us identify only those pixels where the magnetic field is stronger than the threshold.

To ensure the detected regions represent true BMRs and not just unbalanced patches of flux, we apply a flux balance condition at the end of the detection process. This condition ensures that the total positive and negative magnetic flux in a region are reasonably balanced. The flux balance ratio is defined as

$$r = \frac{|\Phi_+| - |\Phi_-|}{|\Phi_+| + |\Phi_-|}, \tag{1}$$

where Φ_+ and Φ_- are the total positive and negative magnetic fluxes in the region, respectively. A region is accepted as a BMR only if the flux balance ratio satisfies $r \leqslant 0.4$. Once the BMRs are identified, we extract the parameters that will be used to analyze this dataset. These parameters are flux-weighted centroid location, latitude (λ), longitude (ϕ), and other relevant parameters such as the maximum field strength in a BMR (B_{max}) and total unsigned magnetic flux (Φ). Finally, the identified region masks are stored for tracking in the next step.

Once detection is complete for all magnetogram data, the stored binary masks undergo preprocessing steps of image opening operation using an adaptive circular kernels to extract the boundaries of the BMR. This modified binary mask and the corresponding magnetograms are used to track the evolution of BMRs. A detailed description of the tracking algorithm can be found in A. Sreedevi et al. (2023); here, we summarize the main steps.

The tracking process begins by isolating a BMR in its precomputed binary mask. Next, we estimate how long this BMR can be observed before it reaches the western limb of the

solar disk. This is done by calculating its heliographic longitude (ϕ) and latitude (θ) and using the known differential rotation profile (R. F. Howard et al. 1990) at its latitude to estimate the time for the BMR to reach the limb. Using this time estimate and the cadence of the observations (96 minutes in our analysis), we determine the maximum number of future observations in which this BMR can potentially be tracked. For each of these future time steps, we predict the expected location of the BMR using the differential rotation profile, followed by applying a feature association technique to find a matching structure.

The feature association algorithm checks that the change in area (i.e., the number of pixels) between two consecutive observations remains within a range of 80%–300%. In cases where data are missing, this condition is made stricter to ensure robust tracking of the same BMR. The tracking continues either until the estimated time of limb passage is reached or until the BMR fails to meet the association criteria described above. Once a BMR is identified and tracked successfully, all its information is stored in a catalog. This catalog includes the unique AutoTAB-ID assigned to each BMR, along with its temporal evolution throughout the tracking period (A. Sreedevi et al. 2023). The current AutoTAB catalog contains information on roughly 12,000 unique BMRs, and the backtracking algorithm is built based on this catalog.

3.2. Backtracking of BMRs Identified by AutoTAB

Upon careful analysis of the BMRs in the AutoTAB catalog, we observed that the AutoTAB missed the initial emergence phase of the BMRs. This is primarily due to AutoTAB's stringent detection criteria, which prioritize decent flux balance between the polarities, a condition usually met at a later stage in the BMR evolution. To address this limitation, we developed a backtracking algorithm that traces each detected BMR backward in time, starting from its first appearance in the AutoTAB catalog, given the limitation of the data associated with the cadence. Our backtracking algorithm is as follows.

Once a BMR is identified by AutoTAB, we extract the region binary mask (saved during the detection) and the corresponding full-disk magnetogram (projection effects corrected) at the time of its first detection (referred as T_0 from now on). The region mask is expanded into a rectangular region by extending 9 pixels ($18''$) outward from the boundaries of the detected region, matching the bounding-box criteria used during the AutoTAB detection phase. We restricted our analysis to BMRs whose first detections occur at heliographic longitudes $\geqslant -35°$, ensuring sufficient preceding observations to trace the region's emergence prior to AutoTAB's initial detection. At the time of first detection (T_0), we calculate three physical quantities within the extracted subregion: (a) $R(T_0)$, the ratio of pixels with absolute magnetic field strength exceeding 100 G to the total number of pixels in the box; (b) $\Phi(T_0)$, the total unsigned magnetic flux of the region; and (c) $B(T_0)$, the flux density, calculated as the unsigned flux divided by the area of the box in pixels. Together, these characterize the magnetic strength and coherence of the BMR at the point of first detection.

Before initiating the backtracking procedure, we apply a preliminary filtering step to remove potentially unreliable detections. Regions with bounding boxes larger than 100 pixels in either width or height are excluded, as such cases are typically indicative of faulty detections carried over from AutoTAB and are unlikely to correspond to genuine BMR emergence. In addition, regions with a $B(T_0)$ (flux density) below a threshold of 15 are discarded, as they often represent weak, incoherent magnetic fields (e.g., salt-and-pepper noise) or decaying remnants of earlier activity rather than newly emerging BMRs. After filtering, the maximum allowable backtracking duration is estimated using the same rotational model adopted in the AutoTAB tracking algorithm.

The algorithm then proceeds by iteratively scanning earlier magnetograms, one by one, in reverse chronological order by differentially rotating the region mask back in time to align with the expected location at previous times. If the mean longitude of the rotated region falls below $-45°$, the procedure is terminated to avoid severe projection effects near the limb. At each valid step, the region's magnetic properties are reevaluated by calculating $R(t)$, the ratio of strong-field pixels (above 100 G) to the total number of pixels in the original box, and $\Phi(t)$, the unsigned magnetic flux. If the region contains no detectable positive or negative polarity pixels, then the step is considered invalid and the algorithm proceeds to the next earlier frame without using the current one. When both polarities are present, two diagnostic ratios are computed to evaluate the evolution of the region relative to its properties at T_0: (a) the pixel ratio $r_p = \frac{R(t)}{R(T_0)}$ and (b) the flux ratio $r_f = \frac{\Phi(t)}{\Phi(T_0)}$. These ratios are used to classify the behavior of the region. If $r_f > 1$ and $r_p > 1$, the region appears to grow unrealistically during backtracking, suggesting an unreliable step. If $r_f \leqslant 0.4$ and $r_p \leqslant 0.5$, the region is interpreted as either decaying rapidly or becoming too weak to be distinguished from the background noise level; in such cases, the step is skipped. If more than five instances are skipped in a row, the backtracking process is terminated. In all other cases, the region is deemed to have successfully backtracked. These thresholds were determined through trial and error and found to be robust against small variations. The final timestamp of backtracking is now the actual emergence time, marked as T_e for the rest of the paper. Backtracking is a very challenging task, as BMRs are not necessarily well flux balanced during the emergence, and the field strength remains near the quiet-Sun (salt-and-pepper) magnetic field value. Despite this, the emergences of some BMRs remain undetected because they either appeared near the east limb, appeared on the Sun's farside, or were already in a decaying phase upon first detection on the nearside. Of 11,987 BMRs initially present in the AutoTAB catalog, we could successfully backtrack 3080 BMRs starting from T_0 to T_e. To show how well our code captures the emerging phases of BMRs, we show the time evolution of two BMRs in Figure 1 (four more cases in Figure 2). As evident in this figure, the code detects the very early emergence phase of these BMRs when they show faint signatures in the magnetogram (and almost no signature in intensity continuum; Figure 6) and effectively tracks them as long as they are on the nearside of the solar disk.

4. Results

After backtracking, we obtain information for 3080 BMRs, covering their evolution from early emergence until either their decay or their passage beyond the west limb, whichever occurs first. The rest of our analysis is based on this catalog. It is worth noting that for some BMRs, the backtracking algorithm
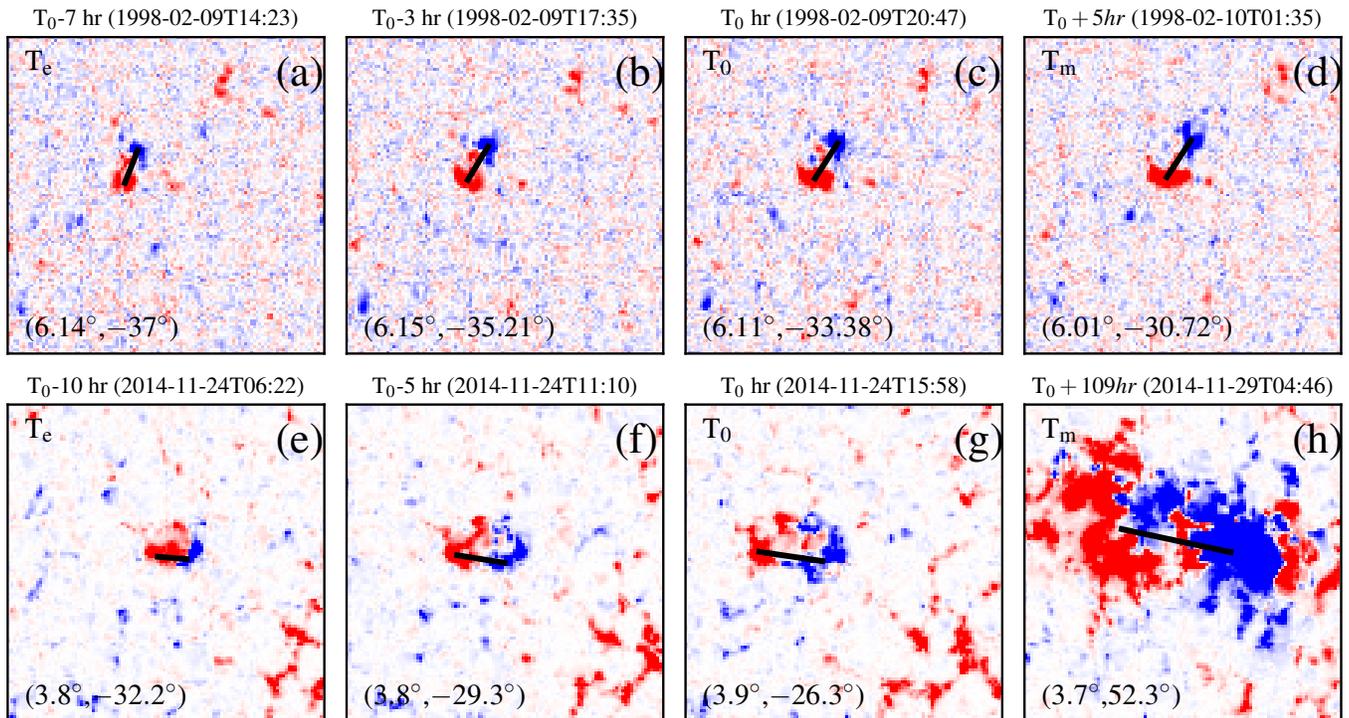
Figure 1. Evolution of two emerging/growing BMRs. (a)–(d) The BMR at the times of initial detection (T_e), at the middle of the backtracking phase, the starting of our backtracking phase (T_0), and at its maximum flux. Note the time sequence: $T_e \rightarrow T_0 \rightarrow T_m$. (e)–(h) The same time frames but for a different BMR. Numbers in brackets on each panel denote the mean latitude and longitude of the region. The mean latitude and longitude of the region are printed in brackets on each panel. LOS field is saturated at 1.5 kG in all panels. The line in each panel connects the flux-weighted centroids of BMR's poles.

identifies T_0 as the same as T_e; such cases are also included in the analysis. Building on this, we now proceed to examine how the statistical properties of these BMRs evolve over time.

4.1. Two Groups of BMRs: Emerging and Discarded BMRs

After a closer look at some of the BMRs in Figures 1 and 2, we see that, in most cases, we can follow BMRs from their earliest stage through to maturity. However, some cases show little or no flux growth over time. In such cases the algorithm fails to capture the emergent phase of the BMRs; instead it captures the decaying phase of the regions or the fragmented parts of an active region. To investigate this, we examined the change in magnetic flux between first detection (T_e) and a later stage (T_0); this helps separate genuine emergences from cases influenced by flux dispersal or reconfiguration of preexisting fields. Figure 3(a) (with an inset on a linear scale) shows a scatter plot of flux at T_e versus flux at T_0. Two clusters are evident: one along the diagonal, indicating little change in flux from the time of emergence to the later time T_0, and another above it, showing the expected growth of emerging BMRs (M. Švanda et al. 2025). The latter fits the conventional expectation of emerging and growing BMRs from the convection zone, while the former does not. Visual inspection confirms that most of the BMRs in the diagonal branch arise from remnant flux of old or decaying regions (Figures 2(a)–(d), fragmented flux from a parent BMR (Figures 2(e)–(h), dispersed or short-lived flux (Figures 2(i)–(p), or ephemeral regions whose (short) emerging phase is missed due to weak flux comparable to the background fields (Figures 2(m)–(p)). Thus, all these are not genuine new emergences.

To objectively separate the two groups, we applied a Gaussian mixture model (GMM) using the angle each point makes with the horizontal axis. This cleanly splits the sample into "emerging" and "discarded" BMRs. The latter group consists of the decaying phases of BMRs or ephemeral regions whose emerging phases are missed or bipolar structures generated from the fragmentation of parent active regions (Figure 3(b)). Only BMRs with ⩾90% probability of belonging to the growing cluster are grouped into the emerging BMR group, highlighted in blue in Figure 3(a), while the red points mark the discarded ones, which make up 27% of the catalog. Comparison of the two groups highlights clear differences (Figures 3(c)–(d)). Growing BMRs typically start with low flux and show steady, rapid increases in unsigned flux (Φ_m) from T_e to T_0, along with consistent growth in footpoint separation. Discarded BMRs often begin with higher flux, which does not increase significantly over the period of their detection because many of them are the decaying BMRs and the ephemeral regions whose fluxes do not grow appreciably during the tracking period. Eventually, their flux distribution remains smaller at T_0 (as seen in Figure 3(a)) or in the matured phase compared to the flux distribution of emerging BMRs. The footpoint separation of these discarded BMRs also shows a distinct evolution. It begins with a large value and increases slowly and irregularly. After manually detecting several of these discarded BMRs in the corresponding intensity continuum data (three are shown in Figure 6), we find that they are not visible as sunspots. All these properties strongly suggest that they arise from remnant or decaying flux rather than fresh emergence from the convection zone, consistent with our case studies. The emerging group/blue points in Figure 3, representing genuine emergences, formed the working sample analyzed separately in A. Sreedevi et al. (2025), where we showed robust Joy's law, starting the time of emergence T_e. In
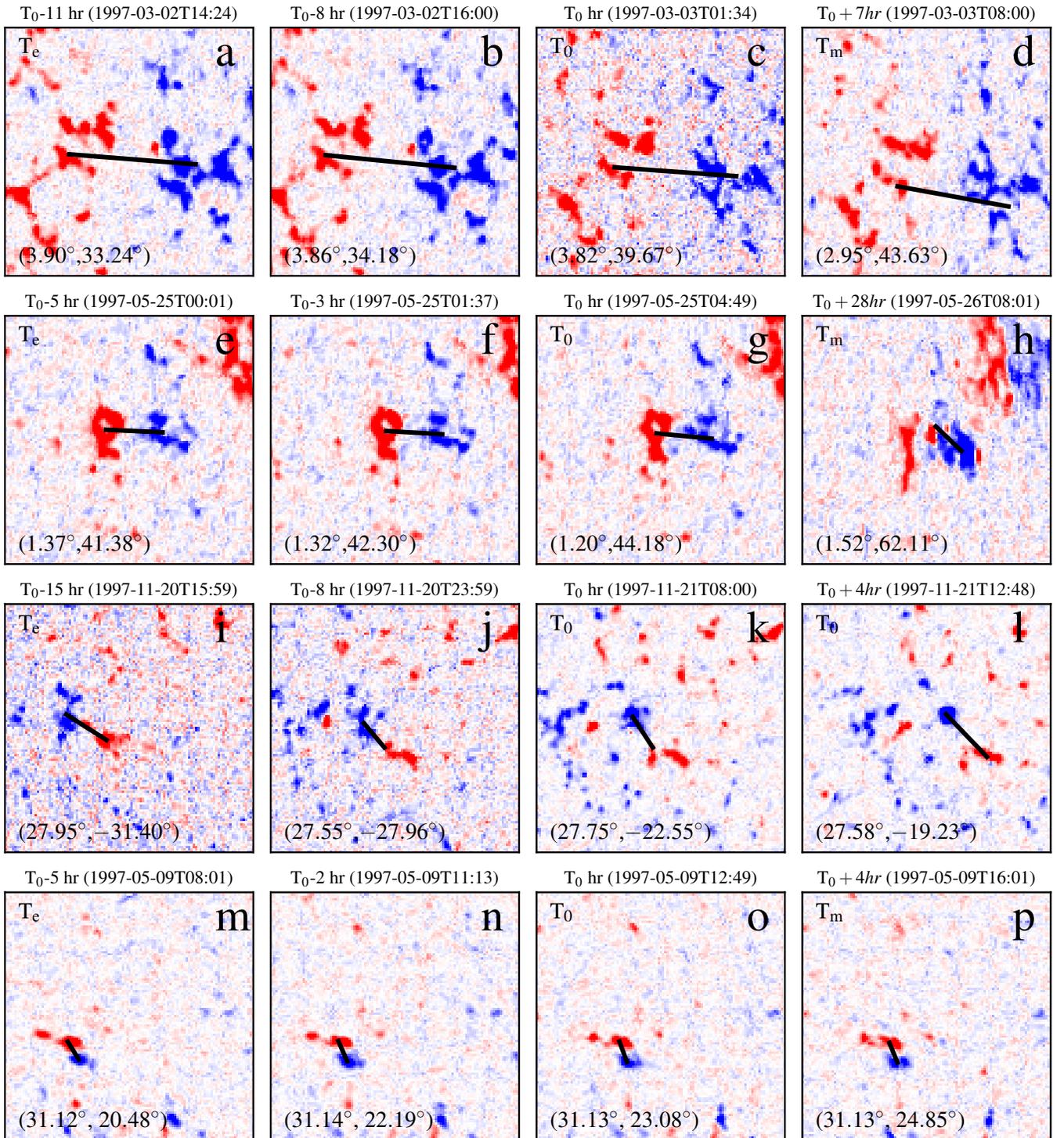
Figure 2. Examples of the time evolution of four discarded BMRs. The figure format is the same as Figure 1, but different rows represent the time evolutions of different BMRs that fall in the same discarded BMR group (red points in Figure 3(a) that were discarded in A. Sreedevi et al. (2025)).

this article, we concentrate on the red/discarded group and how their presence affects the properties of BMRs.

4.2. Tilt Angle of Discarded BMRs

We compute the tilt angle γ relative to the equator employing the usual definition of Y. M. Wang & N. R. Sheeley (1989). In this, BMRs located in the northern hemisphere and exhibiting Hale–Joy orientations are assigned positive tilt

values, and the tilt value is kept within $\pm 90°$. To get a better statistical measure, we also combine the data of two hemispheres by reversing the sign of the tilts of the southern hemisphere.

Figures 4(a), (c), (e) show the tilt distribution of the BMRs that fall into the "discarded" groups. Across all three phases of their evolution—T_e, T_0, and T_m (the matured phase of BMR)— the fitted Gaussian mean stays essentially at $0°$, with no clear sign of systematic tilt. The anti-Hale BMRs are the significant
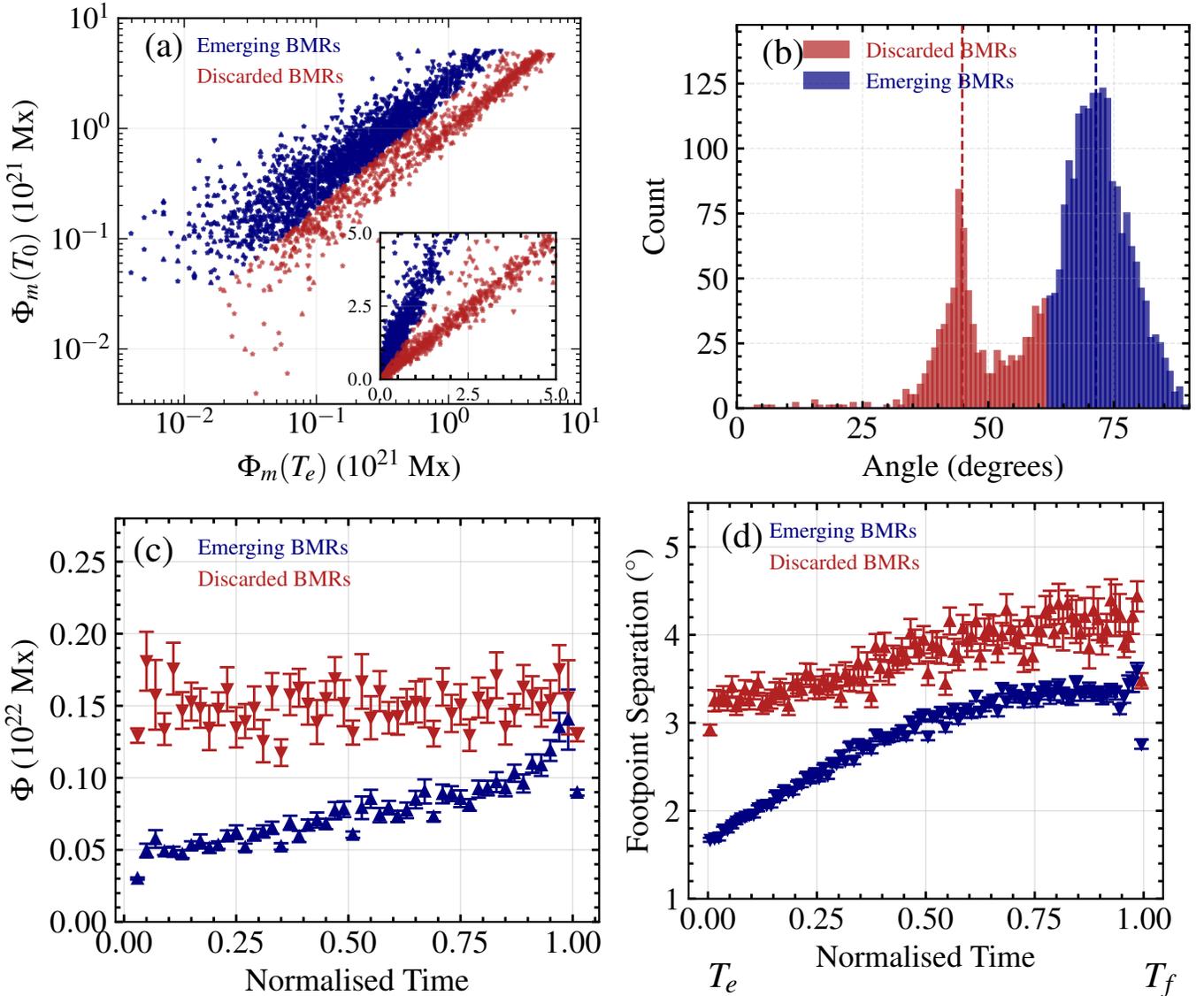
Figure 3. (a) Flux of BMRs (Φ_m) at the start of the backtracking phase (T_0) versus Φ_m at the time of emergence T_e. The inset shows the same but in linear scales. Asterisks, upper triangles, and lower triangles, respectively, represent solar cycles 23, 24, and a part of 25. Red and blue points represent the "discarded" and "emerging" BMRs, isolated based on the angle distribution. (b) Histogram of angles measured at the origin with respect to the horizontal axis. The blue/red Gaussian curves represent the two components obtained from the GMM fit to the angle distribution. Usable BMRs are defined based on the probability of their angles falling into the upper (high-angle) cluster being >90%. (c)–(d) Evolution of the flux and the footpoint separation of "discarded" and "emerging" BMRs as a function of the normalized time with a bin size of 0.01. The time is normalized by the time period between T_0 and T_e for (c) and the total lifetime for (d) of each BMR to bring it within 0 and 1.

contributors in this discarded group. They contribute approximately 26%, 25% and 20% at T_e, $T_{0,}$ and T_m, respectively. These numbers in the case of the emerging BMR group are: 23%, 13%, and 11%, respectively. These anti-Hale (Joy) and Hale-Anti-Joy BMRs, both of which have negative tilts, hinder the development of the polar field, causing variations in the solar cycle and thus posing challenges in solar cycle prediction (A. Biswas et al. 2023; P. Kumar et al. 2024; B. Dey et al. 2025).

When we look at their Joy's law behavior (Figures 4(b), (d), (f)), the picture becomes even clearer: at T_e we see a scattered and even negative γ_0, which gradually shifts toward nearly 0° by T_m, but without any meaningful dependence on latitude. In other words, their tilt evolution is completely inconsistent with the expected Joy's law trend.

This behavior of the tilt angle of discarded BMRs provides an independent check that this group of BMRs does not

comprise genuine emerging BMRs. Therefore, we need to isolate this group from the whole dataset of BMRs identified in our detection code.

4.3. Joy's Law from the Combined Dataset

Since discarded BMRs do not exhibit any characteristic "BMR-like" behavior, in this section we examine how combining the discarded and emerging classes affects the latitude dependence of tilt. Figures 5(a), (c), (e) show the tilt distributions of BMRs at T_e, T_0, and T_m (the mature phase), respectively. At T_e, the fitted Gaussian mean falls at $5°.6 \pm 1°.7$ but with a very wide spread. Over time, the spread (quantified by σ) decreases significantly while the Gaussian mean does not change much as BMRs fully emerge and reach their peak flux. We note that Figure 5(e) should be the same as Figure 2(a) of A. Sreedevi et al. (2024). However, the smaller value of the
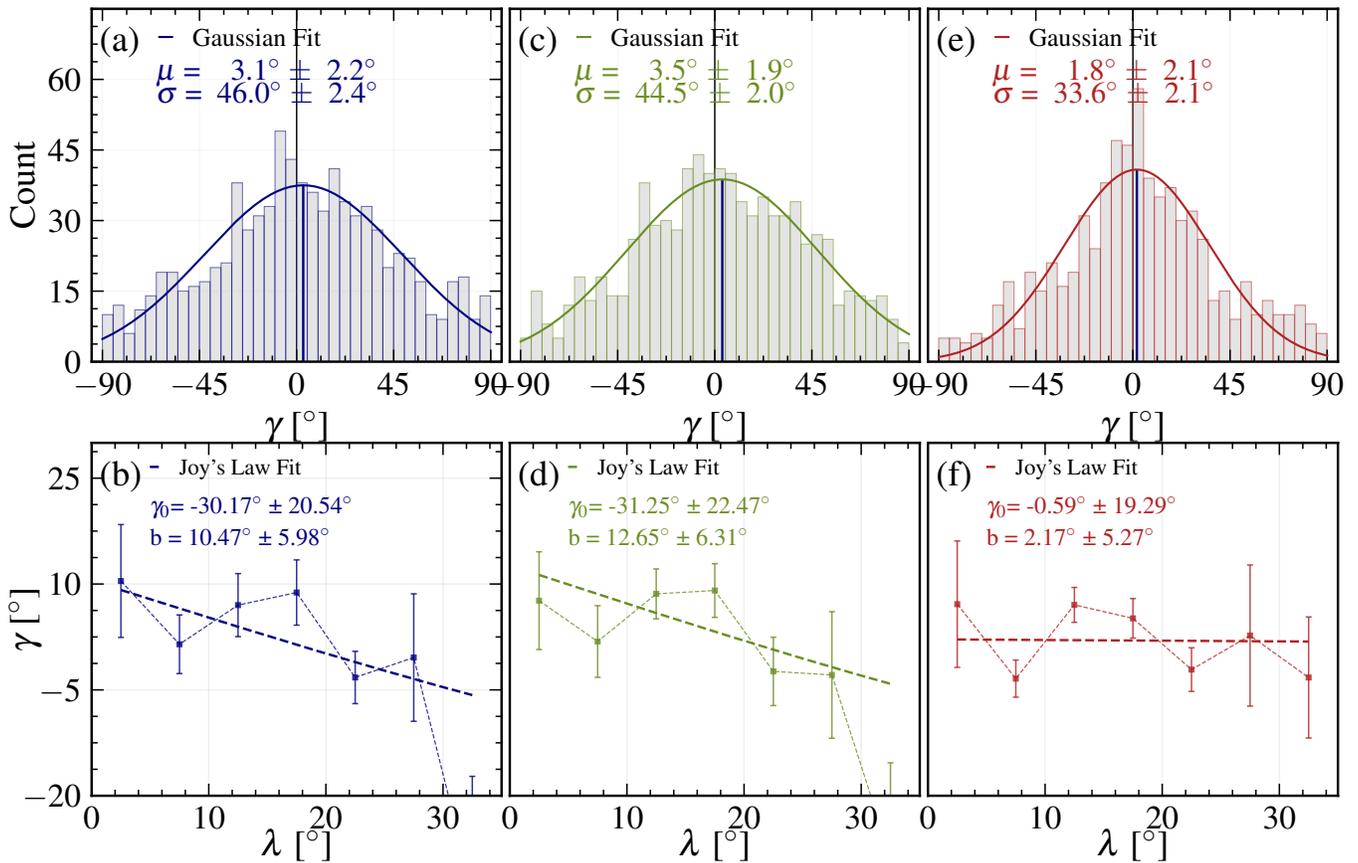
Figure 4. Tilts of discarded BMRs. (a) Distribution and (b) Joy's law of BMR tilt angles at T_e. (c)–(d) At T_0. (e)–(f) At T_m. Joy's law is computed by binning the tilt values in 5° latitude intervals and finding the Gaussian mean tilt for each bin. The error bar is the fitted error of the Gaussian mean.

Gaussian mean in the present data is because only a subset of BMRs could successfully be backtracked in the present study. Many high-flux BMRs that live for longer periods are not backtracked (as they either emerged near the east limb or on the farside of the Sun).

Here in Figures 5(b), (d), (f) we also observe a Joy's law trend, starting from T_e, and the noise around Joy's law decreases steadily as BMRs evolve. However, here the Joy's law trend is quite weaker than we found in our previous study (A. Sreedevi et al. 2025), excluding the discarded BMRs (γ_0 at T_e was 25.98 ± 7.32), demonstrating the influence of the discarded BMRs on the true signature of the BMR. This also iterates the importance of proper identification of the tilts in the region and need of cross validation of tilt in various datasets (T. Baranyi 2015; L. Qin et al. 2025). More specifically, our discarded BMRs isolated simply by observing the flux evolution will be excluded using the cross validation of tilt from white-light measurements as done in L. Qin et al. (2025) because the these discarded BMRs are mostly invisible in white-light observations.

In A. Sreedevi et al. (2025), we show that by separating the discarded BMRs, we find a robust Joy's law at the time of emergence on the solar surface, consistent with the thin flux tube model, which suggests that the Coriolis force causes the tilt of emerging BMRs. If we do not isolate the discarded BMRs from the main dataset as done here, then the scatter around Joy's law is considerably increased, and in a limited dataset, Joy's law may not even be detected. This could be the case in previous studies (A. G. Kosovichev & J. O. Stenflo 2008) where no

evidence of Joy's law was found at the time of emergence of BMRs. Therefore, in any analysis of emerging BMRs, we need to remove the discarded BMRs carefully; otherwise, they may mask the actual properties of real BMRs that the Sun truly produces.

The considerable amount of scatter in the tilt around Joy's law observed at the time of first detection (T_e) originates from two sources. One is due to the inability to correctly detect the very early signature of the BMR and/or incorrect identification of the flux-weighted centroids of two poles, leading to uncertainties in tilt value. This is evident in Figures 7(a)–(d), where a BMR is wrongly identified[5] during the backtracking phase of our automated code and produces incorrect tilts at $t < T_e$. Another source of scatter comes from a physical origin, and it is the turbulent convection. As explained in D. Longcope & A. R. Choudhuri (2002), during the rise of the BMR-forming flux tubes in the convection zone, they are buffeted by convective turbulence in the near-surface layer, and thus their tilts are easily disturbed. However, in the later phase, the Coriolis perturbation overpowers the short-time-scale perturbations of turbulence, which relax more rapidly. Thus, the tilt scatter reduces over the growth of the BMR as observed in our data. The decrease of the scatter in the tilt angle during the growth phase of BMR was also reported in the previous study of H. Schunker et al. (2020) using 153 emerging active regions.

---

[5] By manually observing several BMRs we find that this type of wrong detection during the backtracking phase is quite infrequent.
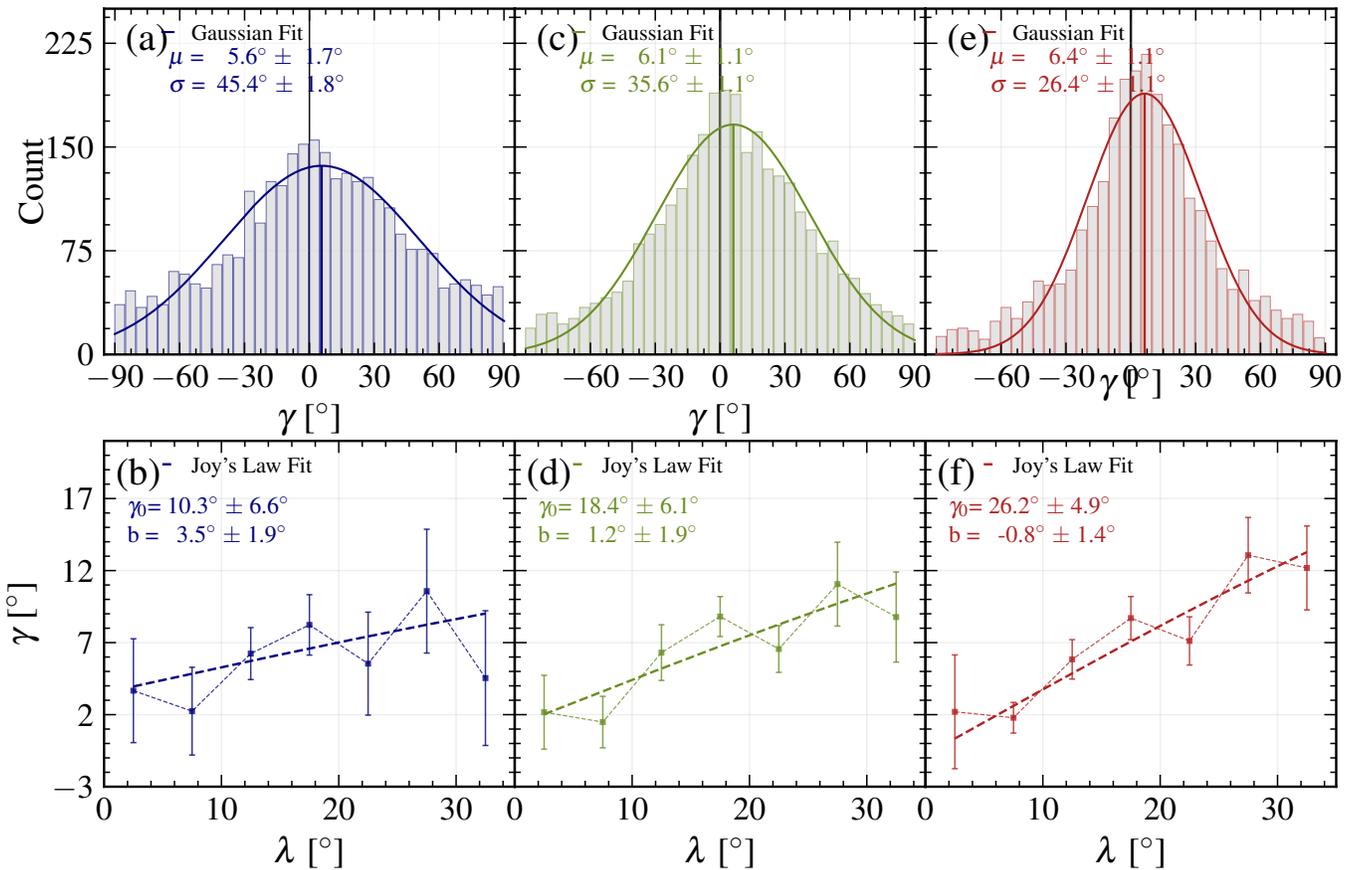
Figure 5. The same as Figure 4 but from all BMRs backtracked in our study, i.e., without excluding the "discarded" BMRs.

5. Conclusion

Various groups have developed algorithms for tracking the BMRs from magnetograms on the nearside, which are used to study their origins and evolutions. AutoTAB is one such state-of-the-art tool. However, existing tracking approaches—whether automated or reliant on human intervention—fall short in identifying the very onset of flux emergence. To address this, we developed a backtracking algorithm that takes the first detection time (T_0) from AutoTAB and traces the BMRs backward to the true moment of polarity appearance, thereby defining a new emergence time (T_e). Applying this method to the 12,000 unique BMRs in the AutoTAB catalog, we were able to backtrack and reliably determine T_e for 3080 cases. This enabled us to systematically investigate their tilt angle distributions and latitude dependence at the moment of true emergence.

By evaluating the flux and footpoint separation evolutions between T_e and T_0, complemented by manual observations of several BMRs, we identified two distinct BMR groups: emerging and discarded. The emerging BMRs exhibit significant flux growth, whereas the discarded ones do not show such considerable growth; instead, they consist of the decaying BMRs with nearby active regions, ephemeral regions, or fragmented bipolar flux originating from a parent active region. We show that the discarded BMRs have almost zero mean tilt and do not show any latitudinal preference like Joy's law, even in their mature phase. The emerging BMRs show nonzero mean tilt and a robust Joy's law even at the time of emergence on the solar surface, suggesting that they are caused by the Coriolis force, as explained by the thin flux tube

model (A. Sreedevi et al. 2025). If we do not isolate the discarded BMRs from the whole dataset obtained from the automated code, then Joy's law trend at the time of emergence weakens as a large scatter accompanies it.

Taken together, these results highlight a crucial implication: including discarded BMRs in statistical studies of solar active regions can severely mask the intrinsic properties of genuine, flux-emerging BMRs. To uncover the actual physical signatures of BMR emergence and their role in shaping Joy's law, it is essential to separate these two populations.

Author Contributions

B.B.K. gave the initial idea. A.S., under the guidance of B.B.K., B.K.J., and D.B., analyzed magnetogram data, developed algorithms for BMR backtracking, and prepared the figures. R.G. analyzed intensity continuum data. A.S. and B.B.K. led the manuscript writing, while B.K.J. commented on the text. All authors contributed to the discussion and presentation of the results.

Appendix

This section presents additional information on the time evolution BMRs in the intensity continuum (Figure 6) and an additional BMR in the magnetogram (Figure 7).
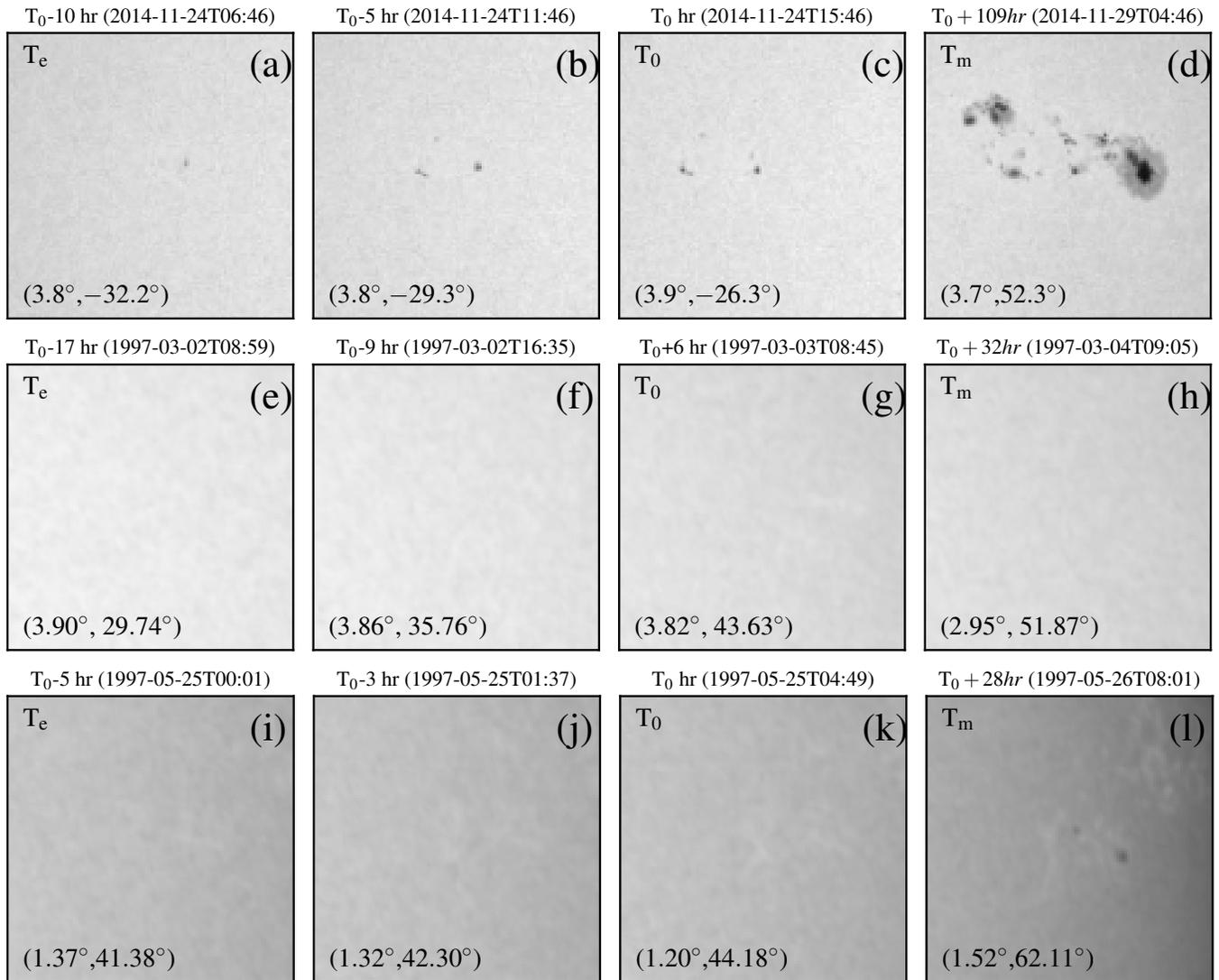
Figure 6. The time evolution of one "emerging BMR" and two "discarded BMRs" whose magnetograms are shown in Figures 1(e)–(h) and Figures 2(a)–(d), (e)–(h), repectively, but are now searched in the corresponding (near-simultaneous) intensity continuum.
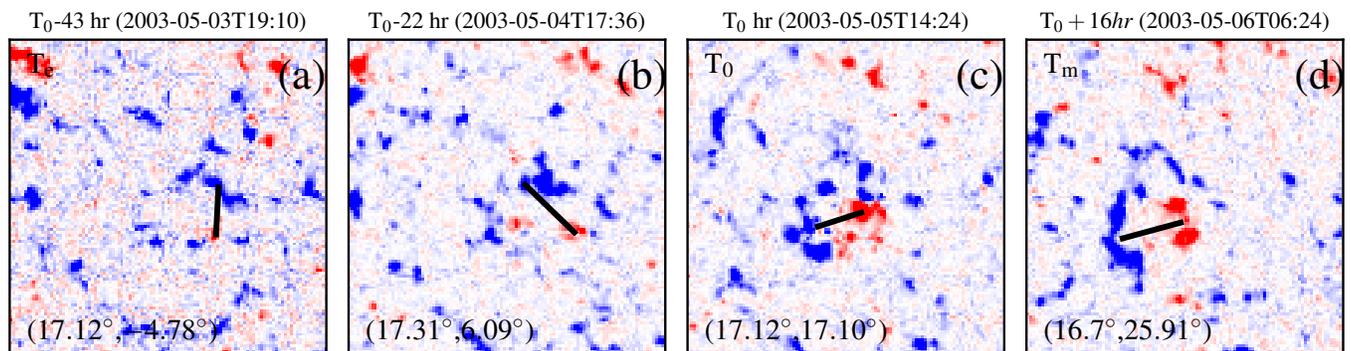
Figure 7. The same as Figure 1 but for a different BMR for which the backtracking phase ($t < T_0$) is wrongly identified.

ORCID iDs

Anu Sreedevi ⬥ https://orcid.org/0000-0001-7036-2902
Bidya Binay Karak ⬥ https://orcid.org/0000-0002-8883-3562
Bibhuti Kumar Jha ⬥ https://orcid.org/0000-0003-3191-4625
Dipankar Banerjee ⬥ https://orcid.org/0000-0003-4653-6823

References

Arlt, R., & Vaquero, J. M. 2020, LRSP, 17, 1
Baranyi, T. 2015, MNRAS, 447, 1857
Biswas, A., Karak, B. B., & Kumar, P. 2023, MNRAS, 526, 3994
Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, SoPh, 289, 3549
Bobra, M. G., Wright, P. J., Sun, X., & Turmon, M. J. 2021, ApJS, 256, 26
Dey, B., Sreedevi, A., & Karak, B. B. 2025, ApJ, 993, 196
Gong, L., Yang, Y., Feng, S., et al. 2024, SoPh, 299, 121
Hale, G. E. 1908, ApJ, 28, 315
Higgins, P. A., Gallagher, P. T., McAteer, R. T. J., & Bloomfield, D. S. 2011, AdSpR, 47, 2105
Howard, R. F. 1991a, SoPh, 132, 49
Howard, R. F. 1991b, SoPh, 136, 251
Howard, R. F., Harvey, J. W., & Forgach, S. 1990, SoPh, 130, 295
Jha, B. K., Karak, B. B., Mandal, S., & Banerjee, D. 2020, ApJL, 889, L19
Karak, B. B. 2023, LRSP, 20, 3
Kosovichev, A. G., & Stenflo, J. O. 2008, ApJL, 688, L115
Kumar, P., Karak, B. B., & Sreedevi, A. 2024, MNRAS, 530, 2895
Longcope, D., & Choudhuri, A. R. 2002, SoPh, 205, 63
Muñoz-Jaramillo, A., Werginz, Z., Vargas-Acosta, J. P., et al. 2016, in 2016 IEEE Int. Conf. on Big Data (IEEE), 3194
Petrovay, K. 2020, LRSP, 17, 2
Qin, L., Jiang, J., & Wang, R. 2025, ApJ, 986, 114
Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, SoPh, 162, 129
Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, SoPh, 275, 207
Schrijver, C. J. 2009, AdSpR, 43, 739
Schunker, H., Baumgartner, C., Birch, A. C., et al. 2020, A&A, 640, A116
Sivaraman, K. R., Gupta, S. S., & Howard, R. F. 1999, SoPh, 189, 69
Sreedevi, A., Jha, B. K., Karak, B. B., & Banerjee, D. 2023, ApJS, 268, 58
Sreedevi, A., Jha, B. K., Karak, B. B., & Banerjee, D. 2024, ApJ, 966, 112
Sreedevi, A., Karak, B. B., Jha, B. K., Gupta, R., & Banerjee, D. 2025, ApJL, 994, L40
Stenflo, J. O., & Kosovichev, A. G. 2012, ApJ, 745, 129
Švanda, M., Jurčák, J., & Schmassmann, M. 2025, A&A, 700, A40
Tlatov, A. G., & Pevtsov, A. A. 2014, SoPh, 289, 1143
Wang, R., Jiang, J., & Luo, Y. 2024, ApJ, 971, 110
Wang, Y. M., & Sheeley, N. R., Jr. 1989, SoPh, 124, 81